Virology 379 (2008) 213–222

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Impact of Δ32-CCR5 heterozygosity on HIV-1 genetic evolution and variability—A study of 4 individuals infected with closely related HIV-1 strains

Kristen Chalmet [a], Filip Van Wanzeele [b], Els Demecheleer [a], Kenny Dauwe [a], Jolanda Pelgrom [b], Bea Van Der Gucht [b], Dirk Vogelaers [b], Jean Plum [a], Lieven Stuyver [c], Linos Vandekerckhove [b], Chris Verhofstede [a,*]

[a] AIDS Reference Laboratory, Ghent University and Ghent University Hospital, De Pintelaan, 185-Block A, B-9000 Gent, Belgium
[b] AIDS Reference Center, Ghent University and Ghent University Hospital, De Pintelaan, 185, B-9000 Gent, Belgium
[c] Virco BVBA, Generaal De Wittelaan, L11 B4, B-2800 Mechelen, Belgium

ARTICLE INFO

Article history:
Received 3 February 2008
Retured to author for revision 8 April 2008
Accepted 30 June 2008
Available online 9 August 2008

Keywords:
HIV-1 evolution
Δ32-CCR5 heterozygosity

ABSTRACT

A cluster of four patients acutely infected with a genetically almost identical virus, allowed us to investigate genetic variability and disease progression in early HIV-1 infection with minimal interference of virus specific factors. Two of the patients were heterozygous for the 32-bp deletion in the CCR5 coreceptor gene. Both showed a slower disease progression with lower viral load levels and a reduced rate of genetic evolution compared to the patients with normal CCR5 alleles. During 3 years of treatment-free follow-up, the mean pairwise genetic distance increased with 1.45% and 1.58% in the two patients with a 32-bp deletion allele compared to 3.05% and 3.57% in the two patients with normal CCR5 alleles. The observed relation between slower disease progression and a reduced evolutionary rate illustrates the influence of the virus replicative capacity, here most possibly hampered by the CCR5 heterozygosity in two of the four individuals, on the genetic evolution of the virus in the host.

© 2008 Elsevier Inc. All rights reserved.

Introduction

In the absence of antiretroviral therapy, the rate of disease progression in HIV-1 infected individuals can be very different, ranging from progression to AIDS within 1 year to long-term asymptomatic survival for over 15 years. It is becoming clear that people with long-term non-progressive infection represent a heterogenous group with non-progression being the result of either infection with a genetically unique virus, an unusually effective immune response, a unique host–virus relation or a combination of these (Deacon et al., 1995; Kirchhoff et al., 1995). The most documented host genetic factor influencing disease progression is a 32-bp deletion from the chemokine CCR5 receptor gene (Husman et al., 1997).

HIV-1 exhibits significant genetic diversity, between hosts as well as within the same individual (Shankarappa et al., 1999). Soon after infection the majority of infected individuals harbour a homogeneous viral population from which a quasispecies evolves over time (Zhu et al., 1993). The ability to generate genetic diversity allows the virus to escape from selection pressures such as the host immune response and antiretroviral drugs and enables the continuous evolution towards the highest degree of fitness. The most variable regions are localized in the viral *env* gene. This gene encodes for the virion surface glycoproteins gp120 and gp41, that are involved in the interaction between the virus and the target cell. Envelope glycoproteins also have an important role in viral fitness (Ball et al., 2003; Marozsan et al., 2001; Rangel et al., 2003) and are known targets for neutralising antibodies (Parren et al., 1999) and $CD8^+$ T cell responses (Carmichael et al., 1996). Patients with slow disease progression are supposed to have vigorous humoral and cellular immune responses that drive the evolution of escape variants (Goulder and Watkins, 2004). Results of several studies indicating a higher degree of virus variation in slower clinical progression support this assumption (Essajee et al., 1999; Ganeshan et al., 1997; Liu et al., 1997; Lukashov et al., 1995; Wolinsky et al., 1996). Others however reported a higher genetic diversity in rapid progressors than in slow progressors (Hutto et al., 1996; Mani et al., 2002; Markham et al., 1998; Strunnikova et al., 1995; Sutthent et al., 1998). The latter support the hypothesis that viral diversity and antigenic variation is mainly driven by viral replication.

The fact that conclusions of a significant number of studies on intra-host HIV variability are so contradictory indicates that the

* Corresponding author. Fax: +32 9 332 38 41.
E-mail addresses: Kristen.Chalmet@ugent.be (K. Chalmet), Filip.Vanwanzeele@uzgent.be (F. Van Wanzeele), Els.Demecheleer@ugent.be (E. Demecheleer), Kenny.Dauwe@ugent.be (K. Dauwe), Jolanda.Pelgrom@uzgent.be (J. Pelgrom), Bea.Vandergucht@ugent.be (B. Van Der Gucht), Dirk.Vogelaers@ugent.be (D. Vogelaers), Jean.Plum@ugent.be (J. Plum), lstuyver@vrcbe.JNJ.com (L. Stuyver), Linos.Vandekerckhove@ugent.be (L. Vandekerckhove), chris.verhofstede@ugent.be (C. Verhofstede).

0042-6822/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2008.06.036

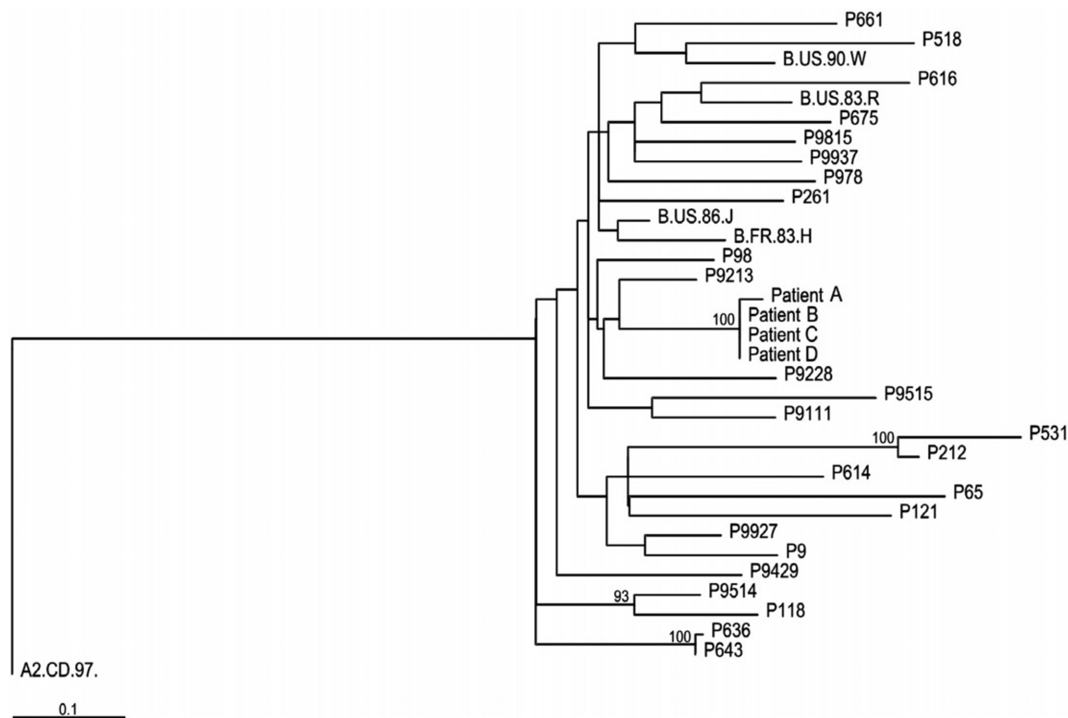

Fig. 1. Phylogenetic tree constructed after analysis of the *env C2–V3* sequence of the 4 study subjects and 25 at random selected individuals from the cohort of patients followed at the hospital. Four subtype B reference strains were included. The subtype A reference strain AF286238 (A2.CD.97) was used as outgroup and root.

relationship between variability and disease progression is complex and still poorly understood. The study of epidemiological-linked HIV-1 infected subjects might serve as the most appropriate model to evaluate the importance of host factors on disease progression and on the HIV variability, because it enables to control for the influence of virus specific factors.

In the current study we have analysed the disease progression and the evolution of HIV-1 quasispecies variability in 4 individuals acutely infected with closely related HIV-1 strains. Three of these individuals were also epidemiologically linked. What makes this clustered cohort of particular interest is the fact that 2 of the 4 subjects are heterozygous for the 32-bp CCR5 deletion. The other 2 patients have two wild type CCR5 alleles. All 4 individuals were temporarily treated with antiretroviral therapy (ART) for one year soon after diagnosis, after which the medication was stopped. Besides a close monitoring of plasma viral load levels and $CD4^+$ T cell counts we performed an extensive analysis of the intrapatient genetic diversity of the HIV-1 *env* gene over a period of 4 years.

Results

Phylogenetic relationship

Phylogenetic analysis of the 300-bp *protease* gene and a 619-bp fragment of the *reverse transcriptase* (RT) gene from 198 patients attending the Aids Reference Centre in our hospital, allowed the identification of a cluster of four patients infected with a closely related HIV-1 strain of the B subtype (results not shown). For three of these patients (B, C and D) a 100% identity of the PR/RT fragment was observed. The viral strain isolated from patient A showed an identity of 99.7%. To further support the genetic relationship, phylogenetic analysis of the *env* gene was performed. The phylogenetic tree shown in Fig. 1 is based on the comparison of the C2–V3 *env* region for the 4 study subjects and for 25 other patients from the hospital cohort with a documented acute subtype B infection as well as 4 subtype B reference strains (B.US.90.WEAU160.U21135, B.US.86.JRFL. U63632, B.US.83.RF.M17451, B.FR.83HXB2-LAI-IIIB-BRU.K03455). The tree was rooted using a subtype A reference strain (AF286238 — A2. CD.97). This analysis confirmed the close genetic identity of the viral strains in the four study subjects.

Viral load and $CD4^+$ T cell count

All four patients received HAART, initiated shortly after their first consultation and continued without interruption for a period of one year. Patients were further followed after treatment

Table 1
Summary of the patients' characteristics

ID	Patient A	Patient B	Patient C	Patient D
CCR5 genotype	CCR5-Δ32/wt	CCR5-wt/wt	CCR5-wt/wt	CCR5-Δ32/wt
HLA type	A01_A02/ B13_B44	A24_A26A/ B07_B35	A02_A68/ B44_B53	A11_A29/ B35_B44
Age at infection	30	31	24	34
Date of the last negative HIV test	Nov '01	Feb '02	Aug '01	Dec '01
Date of the first positive HIV test	Mar '02	Apr '02	Apr '02	Apr '02
Flu-like illness	–	Feb '02	Oct '01	–
Presumed infection date	Before Mar '02	Feb '02	Oct '01	Jan '02
Presumed source of infection	Unknown	Patient C	Unknown	Patient C
Date of inclusion (=day 0)	04/10/2002	05/10/2002	05/14/2002	05/06/2002
Viral load at day 0 (copies/ml)	39,000	424,000	104,000	21,700
CD4 count at day 0 (cells/mm^3)	523	259	361	820
HAART	AZT+ 3TC+ABC	AZT+ 3TC+LPVr	AZT+ 3TC+NFV	AZT+ 3TC+ABC
Time to first undetectable viral load (in days)	112	143	122	55
Period on HAART	04/10/02– 04/24/03	05/10/02– 05/20/03	05/22/02– 05/22/03	05/06/02– 06/07/03
Viral load rebound (1 month after stop)	8820	>100,000	36,100	36,100
Therapy resumed (date)	No	Yes (Oct '06)	Yes (Sept '06)	No

interruption for 3 years. Reinitiating treatment was considered necessary in patients B and C at the end of the follow-up period. The initiative for restarting medication was made by the clinician without knowledge of the findings described in this study. Fig. 2 illustrates the evolution of the viral load and the CD4$^+$ T cell count during the whole follow-up period. Table 1 summarizes the patient characteristics. The two patients with the CCR5-Δ32 allele had lower baseline viral loads and higher baseline CD4$^+$ T cell counts compared to the patients with wild type CCR5. The viral load decreased to below detection limits within 55 to 143 days after treatment initiation in all 4 patients. The mean viral load rebound at one month after treatment interruption was lower for the two CCR5-Δ32 heterozygotes compared to the patients with wild type CCR5 (8820 copies/ml and 36,100 copies/ml versus 36,100 and >100,000 copies/ml) but the difference was not significant (p=0,7). After viral rebound, the two CCR5-Δ32 patients were able to suppress viral replication to a set-point level below this of the two patients with wild type CCR5 alleles (see Fig. 2). The low viral load levels were maintained over the whole 3 years of treatment-free follow-up. In the CCR5-Δ32/wt patient D results of less than 1000 RNA copies/ml were obtained for most of the measurements. The same patient was also able to conserve his CD4$^+$ T cells (lowest CD4$^+$ T cell count: 800 cells/mm^3). One of the patients with a wt/wt CCR5 gene (patient B), showed a fast decrease of the CD4$^+$ T cell count after treatment interruption to values below 350 cells/mm^3. In patients A (CCR5-Δ32/wt) and C (CCR5-wt/wt) the CD4$^+$ T cell counts fluctuated between 500 and 800 cells/mm^3.

Phylogenetic analysis of the envelope gene

In order to characterize the genetic diversity and the genetic evolution of the virus in the different hosts, single genome sequencing of the V1–V4 *env* region from PBMC-derived viral DNA was performed. A relatively large *env* fragment of >1000 nucleotides was sequenced from 4 to 5 samples of each patient, collected between the first visit (day 0), and 4 years thereafter (3 years after treatment interruption). At least 10 individual sequences (between 10 and 19) were analysed per sample. Limiting dilution PBMC-derived DNA sequences were obtained for all samples collected. Additionally, clonal RNA sequences were obtained for the plasma samples collected at baseline and 3 years after treatment interruption. An overall consensus sequence was composed from all baseline sequences (DNA and RNA) and used to root the phylogenetic trees (Fig. 3). The trees in Fig. 3 illustrate the tight intermingling of the baseline sequences from patients B, C and D. Baseline samples from patient A cluster away (Fig. 3A). The sequences of the samples collected during the last visit form separate clusters for all 4 patients, indicating a divergent evolution (Fig. 3B). Differences in evolutionary rate for the 4 patients are apparent from the tree topology. The trees also show the close intermingling of DNA and RNA sequences for each of the 4 patients.

Analysis of the intra-sample sequence heterogeneity

The intra-sample virus heterogeneity was assessed by calculating the mean pairwise genetic distance between the clonal sequences of the

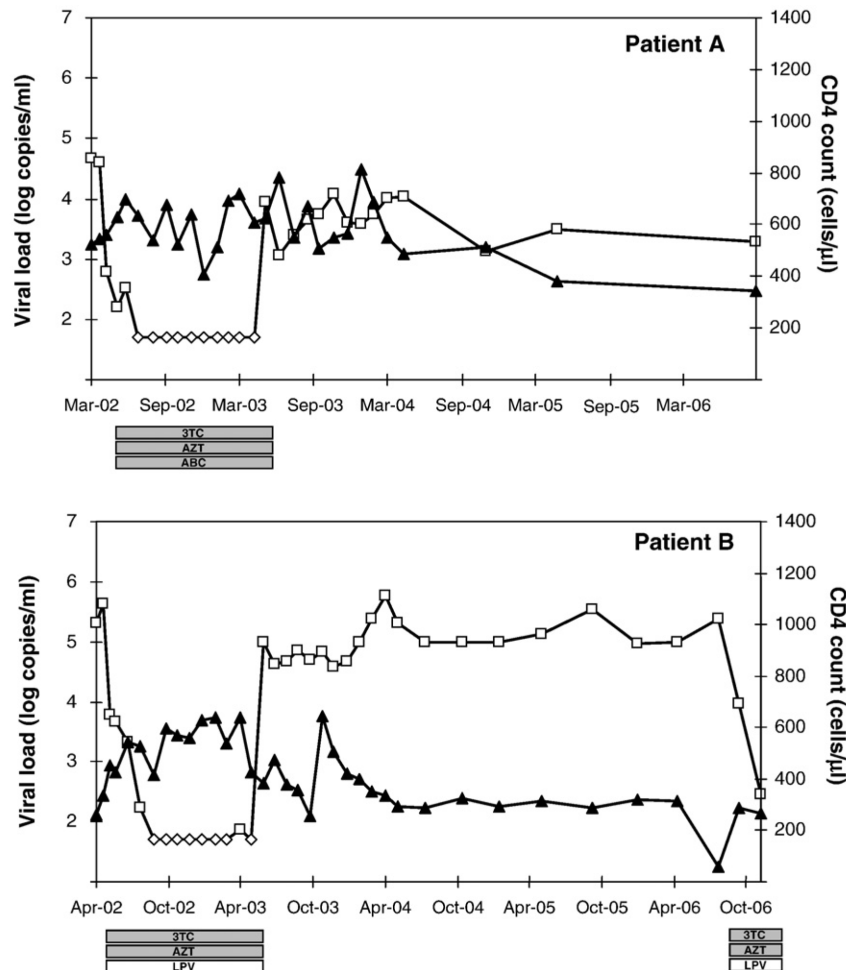

Fig. 2. Evolution of the viral load (□—) and CD4$^+$ T cell count (▲ --) over time in the 4 study subjects. HAART was initiated for 1 year starting soon after collection of the first sample. Viral load remained undetectable (<50 copies/ml) during the treatment period in all 4 patients. The CD4$^+$ T cell count is expressed as cells/mm^3. The viral load is expressed as log copies/ml. ◊: viral load of <50 copies/ml (log copies/ml <1.70). 3TC = lamivudine; AZT = zidovudine; ABC = abacavir; LPV = lopinavir; NFV = Nelfinavir.

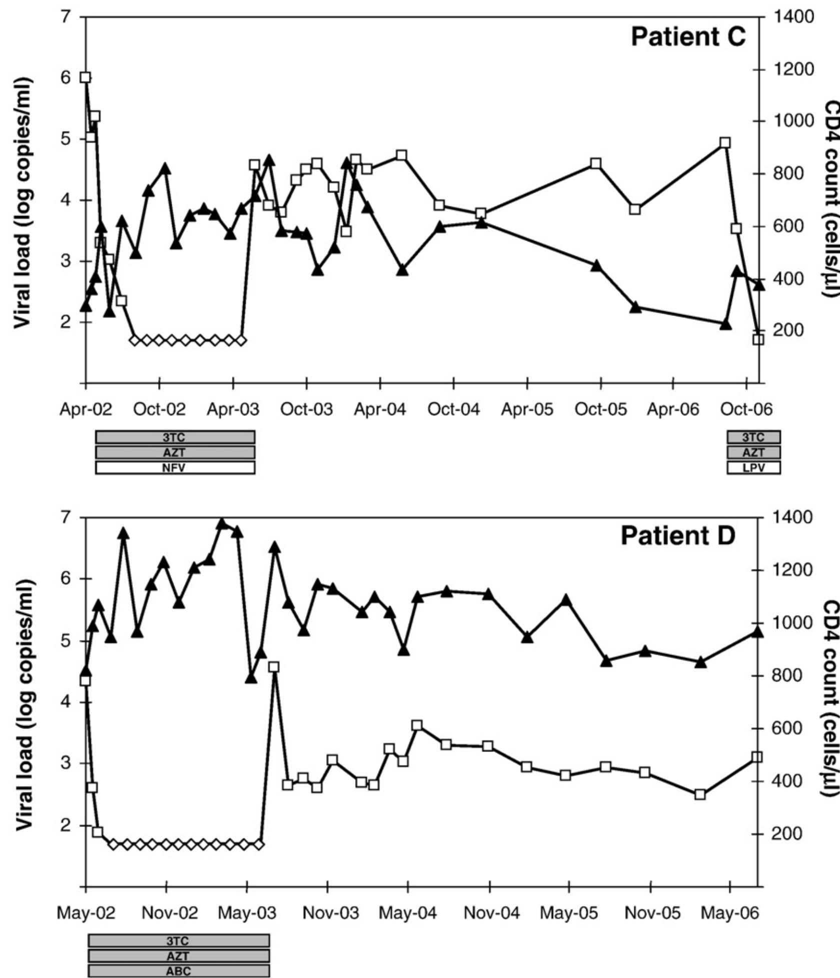

Fig. 2 (*continued*).

same sample. The results reveal a low intra-sample heterogeneity in the baseline samples and in the samples collected one month after treatment interruption (Table 2). The baseline heterogeneity was the highest in patient A. Patient A was also the only patient with a slight increase in heterogeneity between baseline and treatment stop. In the other patients the heterogeneity was stable or decreased during the one-year treatment period. After treatment interruption, the mean pairwise genetic distance in PBMC-derived viral DNA increased with 1.58% (0.48%/yr) and 1.45% (0.49%/yr) in the two patients with a 32-bp CCR5 deletion (patients A and D) and with 3.05% (0.97%/yr) and 3.57% (1.26%/yr) in the two patients with normal CCR5 alleles (patients C and B). The increase in virus heterogeneity correlates with the set-point viral load. The intrapatient virus heterogeneity in the last sample, collected 3 years after treatment interruption, was the highest for patient B followed by patient C, D and A (Table 2). The pairwise comparisons of the differences in virus heterogeneity between patients were statistically significant for all but patients B and C ($p=0.203$). With exception of the baseline sample for patient B, the genetic heterogeneity observed for the RNA and DNA sequences was identical.

Apart from differences as a result of point mutations, an increasing number of insertions and deletions were observed over time, especially in the V1–V2 region. In order to account for these insertions and deletions as an additional contribution to the overall variability in the *env* region, the mean fragment length and the range of the fragment lengths for each sample was calculated (Table 2). Patient D was the only patient in whom the total fragment length of the different sequences remained almost constant. In the other patients a clear trend towards increased fragment length was observed.

Analysis of the genetic divergence

The genetic evolution in the individual patients over time was assessed by calculating the mean pairwise genetic distance between the sequences of the baseline sample and each of the follow-up samples (Table 2). The evolution of the mean pairwise genetic distances from baseline are illustrated in Fig. 4. Patient C is evolving with the highest rate, followed by patient B, A and D. Differences in the slopes of the regression line are statistically significant for the comparison between patients A and C, between patients B and D and between patients C and D. There was a trend for a steeper curve for patient B compared to patient A but this difference did not reach statistical significance ($p=0.074$). No differences in the shape of the regression lines were observed for the two patients with a mutant CCR5 (patients A and D; $p=0.66$) and for the two patients with wild type CCR5 (patients B and C; $p=0.20$).

Discussion

This paper describes a cluster of 4 individuals acutely infected with a highly related HIV-1 strain. Two of the subjects were shown to be heterozygous for the 32-bp deletion in the CCR5 gene. The recent onset of infection in all 4 was confirmed by the observation of a homogeneous virus population in the first sample collected. Baseline *env* sequences were tightly intermingled. Though it is impossible to exclude that the founder virus for patients A and D was 'less fit' than the founder virus for patients B and C this possibility is considered highly unlikely considering the close phylogenetic relation between

Table 2
Follow-up data on viral load, CD4 count, virus genetic evolution and virus heterogeneity

Patient ID	CCR5 genotype	Collection date	Day	Sample	Viral load	CD4 count	*N*[a]	Mean fragment length[b] (range)	MPD from baseline[c]	SD	MPD within sample[d]	SD
A	Δ32	4/10/2002	0	DNA	39,000	523	14	1011.00 (1011–1011)	–	–	0.82	0.47
A		**4/10/2002**	**0**	**RNA**	**39,000**	**523**	**43**	**1011.00 (1011–1011)**	**–**	**–**	**0.74**	**0.43**
A		5/22/2003	402	DNA	8820	621	12	1010.75 (1008–1011)	0.99	0.47	1.07	0.46
A		4/28/2004	738	DNA	10,700	486	13	1007.08 (993–1011)	1.49	0.62	1.13	0.54
A		5/11/2005	1111	DNA	2990	378	11	1003.91 (987–1011)	1.40	0.63	1.54	0.71
A		9/13/2006	1593	DNA	1960	345	10	1016.10 (993–1062)	3.44	0.92	2.66	0.78
A		**9/13/2006**	**1593**	**RNA**	**1960**	**345**	**44**	**1021.09 (1008–1041)**	**–**	**–**	**2.80**	**1.36**
B	WT	5/10/2002	0	DNA	424,000	259	12	1011.00 (1011–1011)	–	–	0.39	0.19
B		**5/10/202**	**0**	**RNA**	**424,000**	**259**	**38**	**1011.00 (1011–1011)**	**–**	**–**	**0.70**	**0.27**
B		6/20/2003	400	DNA	>100,000	385	10	1011.00 (1011–1011)	0.37	0.18	0.20	0.13
B		11/17/203	547	DNA	68,300	646	10	1006.50 (999–1011)	0.64	0.25	0.91	0.34
B		5/26/2004	736	DNA	210,000	294	12	1009.50 (996–1014)	1.34	0.54	1.27	0.33
B		5/24/2005	1094	DNA	131,000	314	11	1024.36 (1011–1047)	2.64	1.09	2.68	0.89
B		5/3/2006	1433	DNA	>100,000	315	10	1044.90 (1026–1053)	4.21	0.44	3.96	1.90
B		**5/3/2006**	**1433**	**RNA**	**>100,000**	**315**	**32**	**1044.75 (1005–1053)**	**–**	**–**	**3.94**	**2.47**
C	WT	5/14/2002	0	DNA	104,000	361	13	1009.39 (1008–1011)	–	–	0.65	0.23
C		**5/14/2002**	**0**	**RNA**	**104,000**	**361**		**1009.87 (1008–1011)**	**–**	**–**	**0.73**	**0.28**
C		6/23/2003	399	DNA	36,100	716	10	1011.00 (1011–1011)	0.61	0.27	0.45	0.28
C		11/28/2003	554	DNA	39,300	437	14	1010.79 (1008–1011)	1.22	0.40	1.01	0.44
C		3/31/2004	677	DNA	31,800	675	13	1024.62 (1011–1044)	1.93	0.59	1.64	0.53
C		9/24/2004	850	DNA	8200	600	14	1035.00 (1011–1053)	2.38	1.04	2.50	1.02
C		9/5/2006	1551	DNA	85,100	227	13	1043.54 (1032–1065)	6.42	1.27	3.50	1.90
C		**9/5/2006**	**1551**	**RNA**	**85,100**	**227**	**45**	**1045.32 (1029–1065)**	**–**	**–**	**3.18**	**1.48**
D	Δ32	5/6/2002	0	DNA	21,700	820	17	1011.00 (1011–1011)	–	–	0.20	0.14
D		**5/6/2002**	**0**	**RNA**	**21,700**	**820**		**1011.00 (1011–1011)**	**–**	**–**	**0.30**	**0.17**
D		7/7/2003	421	DNA	36,100	1290	11	1011.00 (1011–1011)	0.16	0.12	0.13	0.09
D		3/26/2004	680	DNA	1710	1040	16	1011.00 (1011–1011)	0.41	0.28	0.54	0.30
D		8/2/2004	806	DNA	2050	1120	11	1011.27 (1011–1014)	0.78	0.25	0.64	0.23
D		2/1/2005	985	DNA	868	950	15	1011.00 (1011–1011)	0.84	0.37	0.72	0.22
D		**2/28/2006**	**1372**	**RNA**	**315**	**851**	**40**	**1011.00 (1011–1011)**	**–**	**–**	**0.99**	**1.08**
D		7/6/2006	1500	DNA	1240	969	18	1011.00 (1011–1011)	1.48	0.95	1.58	0.62

The viral load is expressed as copies/ml, the $CD4^+$ T cell count is expressed as cells/mm^3.
RNA samples are indicated in bold.
[a] number of single cell or clonal sequences used for the analysis.
[b] mean total nucleotide length of the fragments.
[c] % mean pairwise difference (MPD) between the baseline sequences (day 0) and the sequences of each of the follow-up samples.
[d] % intra-sample mean pairwise distance (MPD).

the viruses in all 4 individuals and the fact that the transmissions in this cluster occurred during the early stage of infection at a time when a relatively homogeneous virus population was present. The aim of this study was to follow the disease process and to analyse the genetic evolution of the virus over a 4-year time period as a way to assess the influences of host genetics and more specifically the influence of the CCR5 genotype on these processes.

The results of several studies indicate the importance of the host genetic background in the HIV disease process (Draenert et al., 2006; Hutto et al., 1996; Liu et al., 1997; Mikhail et al., 2005). Probably one of the most intensively studied host–virus relations is the influence of the CCR5 genotype. Analysis on large databases of individual patients has shown that CCR5-Δ32/wt heterozygotes have a reduced risk for progression to AIDS and lower viral loads, especially early in the course of their disease (Ioannidis et al., 2001). The observations made in the small clustered cohort presented here, confirm these findings. But, despite infection with an almost identical virus and the same CCR5 genotype, individual differences in both disease progression and viral load were observed, pointing to the possible influence of other host factors. Our cohort was too small to study other genetic factors such as the HLA type or to assess the influence of other STD infections.

No correlation between the viral load and virus heterogeneity was observed in the first samples collected from the patients although the slowest progressor, patient D, already presented with the lowest viral load and the most homogeneous virus population at baseline. Comparison of the heterogeneity of the virus population before ART and 1 month after ART interruption, showed a restriction of the genetic evolution during treatment in 3 of the 4 patients with a slight increase in patient A. A stabilisation of the genetic diversification during HAART initiated in a later stage of infection has been reported (Troyer et al., 2005) but ongoing low level viral replication in the presence of HAART can results in a further accumulation of mutations (Gunthard et al., 1999) and might explain the increase in heterogeneity observed in patient A. After treatment interruption, the virus population steadily evolves to a heterogeneous mixture, though with a different rate from patient to patient. The divergence from the baseline sample over the 4 years of follow-up was not significantly different for the patients with the same CCR5 genotype but a faster rate of divergence was observed for the two CCR5-wt/wt patients compared to the CCR5-Δ32/wt. Also the intra-sample virus heterogeneity determined in the last sample collected was significantly higher in the patients with CCR5-wt/wt than in the CCR5-Δ32/wt patients.

The reduced rate of viral replication, evolution and immune deterioration was the most apparent for the CCR5-Δ32/wt patient D. The CCR5-Δ32/wt patient A seems to evolve faster, though still at a lower rate than the CCR5-wt/wt patients B and C. The higher virus heterogeneity in the baseline sample of patient A compared to the others and the clustering away of the sequences from patient A in the baseline phylogenetic tree, might indicate that patient A was infected a few months longer than the other patients at the time of his first presentation and this could have influences the results obtained for genetic evolution and intra-sample heterogeneity. However, the observed differences in virus replication and genetic evolution between the two CCR5-Δ32/wt patients as well as between the two CCR5-wt/wt patients can also be seen as an indication for the interference of other host factors.

In adults, substitutions in a partial *env* fragment are supposed to accumulate at a rate of approximately 1% per year (Shankarappa et al., 1999). Taken into account a delay in virus evolution during the one year HAART, the evolutionary rates obtained for the patients with the

Fig. 3. Phylogenetic trees for the V1–V4 *env* sequences obtained for the samples collected at the first visit (baseline) (A) and 3 years after the HAART interruption (B). The bar denotes 1% nucleotide diversity for panel A and B. ◆: Patient A RNA sequences; ◊: Patient A DNA sequences; ■: Patient B RNA sequences; □: Patient B DNA sequences; ▲: Patient C RNA sequences; △: Patient C DNA sequences; ●: Patient D RNA sequences; ○: Patient D DNA sequences.

Fig. 3 (*continued*).

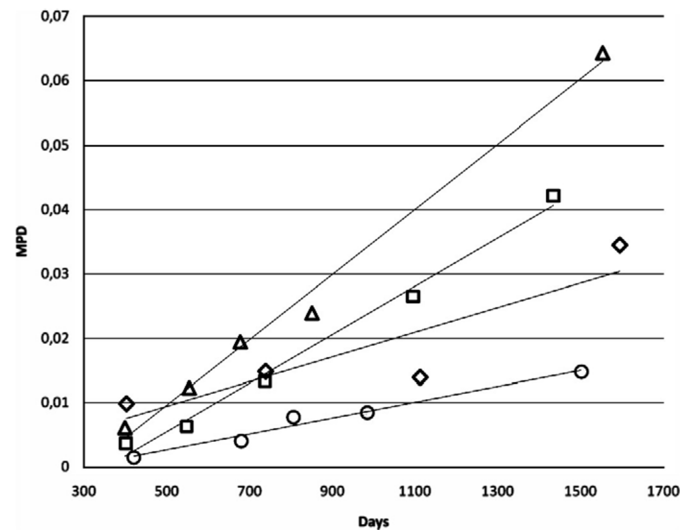

Fig. 4. Genetic evolution after treatment interruption. The *X*-axis represents the mean pairwise genetic distances between baseline sequences and the sequences for each time point after treatment interruption. The *Y*-axis shows the collection day, with day 0 = baseline. ◊ Patient A; □ Patient B; △ Patient C";; ○ Patient D.

wt/wt CCR5 alleles, B and C, are in agreement with these reports (resp. 1.26%/yr and 0.97%/yr). A much slower evolutionary rate (resp 0.48%/yr and 0.49%/yr) was seen in the individuals with a 32-bp deletion in the CCR5 gene, patients A and D, correlating with the low viral load. Although the relation between low genetic diversity and low viral load or slow clinical progression is generally accepted (Mani et al., 2002; Markham et al., 1998; Mikhail et al., 2005; Sutthent et al., 1998), some reports have described an inverse relationship between viral load or clinical progression and genetic diversity (Essajee et al., 1999; Ganeshan et al., 1997; Liu et al., 1997; Wolinsky et al., 1996).

Besides point mutations, extensive insertions and deletions were observed in the *env* gene, especially in the V1–V2 region. A remarkable observation is the increase of the *env* fragment length over time in 3 of the 4 patients. Only the non-progressing patient D conserved the original fragment length. The meaning of this observation is unclear.

The results of this study also revealed a comparable genetic diversity and a close intermingling of the PBMC-derived DNA and the plasma derived RNA sequences. This finding indicates that PBMC-derived viral DNA can adequately reflect the potential of replicative competent viruses that is available and proves that our results are not influenced by an overrepresentation of defective sequences.

All 4 subjects received ART for one year started within 6 months of acute infection. Although there was a treatment bias, with the two CCR5-Δ32/wt patients receiving a NRTIs based regimen (3TC+AZT+ABC) and the two CCR5-wt/wt patients receiving a PI based regimen (3TC+AZT+LPV/NFV), there are no reasons to believe that these differences influenced the further disease course, though we cannot exclude that the HAART has further reduced the already lower rate of viral evolution in individuals with CCR5 delta-32 mutation. At the time of initiation, early ART was supposed to allow preservation of the HIV-1 specific proliferative response and improve the control of the virus after treatment interruption. Today only few studies have confirmed this hypothesis (Rosenberg et al., 2000).

In conclusion, the results of this study show the different rate of genetic evolution of a nearly identical virus in different hosts and demonstrate a correlation between genetic evolution, viral load and CCR5 genotype. Two individuals with a CCR5-Δ32/wt genotype had less HIV-1 viremia and a slower genetic evolution of the virus compared to the 2 individuals with a CCR5-wt/wt genotype. These results are in agreement with the hypothesis that the viral replicative capacity has an important influence on virus variability, at least in the early stages of infection. But individual differences in viremia, genetic variability and fluctuations of $CD4^+$ T cell numbers, irrespective of the CCR5 genotype, point to the additional influence of other host factors in the disease process.

Materials and methods

Study subjects

Four HIV-1 infected individuals, from the Aids Reference Centre (ARC) of the Ghent University Hospital, Belgium, were recognized as being infected with closely related HIV-1 strains after phylogenetic analysis of the reverse transcriptase and *protease* gene. All 4 subjects were infected through homosexual contact. They all had a regular HIV screening performed allowing a very precise estimation of the infection date. Two of them experienced a period of influenza-like illness, preceding a positive HIV test. Subject C was the source of infection for both subjects B and D. Transmission occurred at a time when subject C was himself acutely infected but still unaware of his HIV status.

From patient A it was known that he had a sexual contact with subject B but according to his information this contact dated from a time when subject B was still HIV negative. Subject A had frequent anonymous contacts and tracing the real source of his infection was therefore difficult.

All four patients were acutely infected at the time of their first consultation and a first sample was drawn within six months of the presumed infection date. All four individuals participated in a clinical study aimed at evaluating the immunological and virological benefits of a one-year course of ART initiated within 6 months after infection.

Patient C suffered from other sexually transmitted infections (STD) during the follow-up period (Syphilis, HBV, HCV, Lymphogranuloma venereum).

The patients signed an informed consent form for participation in this study and the project was approved by the Ethical Committee of the institution.

Samples

EDTA blood samples were collected regularly. The plasma fraction was stored at −70 °C. Peripheral blood mononuclear cells (PBMC) were recovered by centrifugation on a Ficoll-Hypaque gradient and either stored at −70 °C or cryopreserved in liquid nitrogen. Phylogenetic analysis was performed on PBMC-derived DNA sequences of samples collected at different time points between the initial presentation in

the clinic (day 0) and 4 years thereafter (3 years after treatment interruption) and on plasma derived RNA sequences from blood samples collected at day 0 and at day 1593, 1433, 1551 and 1372 for patients A, B, C and D respectively.

Quantification of RNA plasma viral load

Plasma HIV-1 RNA was quantified with the Amplicor HIV Monitor test kit v1.5 (Roche Diagnostics Systems, Basel, Switzerland). The lower limit of detection of this assay is 50 RNA copies/ml. The $CD4^{+}$ T cell count was performed by flow cytometry, using the FACScan cytofluorometer and the Cellquest software (Beckton Dickinson Mountain View, California, USA), on freshly drawn blood samples. Absolute $CD4^{+}$ T cell numbers were expressed per microliter.

Detection of the CCR-5 deletion

The presence of a 32-bp deletion in the CCR-5 gene was assessed by amplifying a 239-bp fragment flanking the described 32-nucleotide deletion in the CCR-5 gene with primers 5′-GATAGGTACCTGG-CTGTCGTCCAT-3′ (sense, position 612 to 635) and 5′-AGATAGTCATCTT-GGGGCTGGT-3′ (antisense, position 829 to 850) as described (Husman et al., 1997). PCR products were analysed on a 3.5% agarose gel. To confirm the presence of the deletion, the bands in the agarose gel were sliced out and a sequencing reaction was performed using the same primers as the ones used for PCR amplification.

Population envelope sequencing of plasma viral RNA

Viral RNA was extracted from 500 µl plasma with the High Pure nucleic acid kit (Roche Molecular Biochemicals, Mannheim, Germany). An outer *env* RT-PCR product was obtained using the Titan One tube RT-PCR system (Roche Molecular Biochemicals, Mannheim, Germany) with 0.1 µM of the forward primer (5′-GAGGATATAATCAGTTTATGG-3′) and the reverse primer (5′-GGTGGGTGCTATTCCTAATGG-3′). Two µl of the outer RT-PCR product was subsequently used for a nested amplification with the primers 5′-GATCAAAGCCTAAAGCCATG-3′ and 5′-ACTTCTC-CAATTGTCCCTCATAT-3. The second-round PCR was carried out using Taq polymerase (Applied Biosystems Incorporated, Foster City, CA). Direct sequencing of both sense and antisense strands was done with the BigDye® Terminator Cycle Sequencing kit v. 3.1 (Applied Biosystems, Foster City, California, USA) using the inner amplification primers and two internal primers: 5′-AATTCCATGTGTACATTGTACTG-3′ and 5′-TATTGTGCYCCRGCTGGTTTTGC-3′. Sequencing reaction products were analysed on an ABI 310 Genetic Analyser (Applied Biosystems). A relatively conserved fragment comprising the C2 and V3 region (300 bp) was cut out of the obtained sequences, aligned manually and used for phylogenetic analysis of the population *env* sequences from different subtype B infected patients from the clinic.

DNA extraction, limiting dilution PCR and single genome sequencing

DNA was extracted from thawed PBMC using the QIAamp Blood Kit (QIAGEN GmbH, Hilden, Germany). After extraction, all DNA samples were diluted tenfold. Nested-PCR amplification of a fragment spanning the V1 to V4 region of the *env* gene (nucleotide positions 6591 to 7607 in the HIV-HXB2 genome) was performed using the same primers as described for the population RNA sequencing. The lower limit of detection of the PCR assay was equivalent to 1 copy per reaction.

Approximately 40 identical PCR reactions were run for each tenfold diluted sample. If less than two-third of those reactions were found positive, the PCR products were selected for sequence analysis. Both positive and negative controls were included in all PCR assays to assess the sensitivity of the reaction and to detect possible contamination. PCR and sequencing reactions were performed for one sample at a time to avoid mixing up of samples from different time periods or from different subjects. Direct sequencing of both sense and antisense strands was done as described above with the BigDye® Terminator Cycle Sequencing kit v. 3.1 (Applied Biosystems, Foster City, California, USA). All sequencing products for which inspection of the electropherograms revealed nucleotide-mixtures at one or more positions were withdrawn from further analysis. At least 10 single cell PCR products were sequenced per sample. The whole sequenced fragment (approximately 1014 bp) was used for phylogenetic analysis.

RNA extraction, amplification and clonal sequencing

RNA was extracted from plasma using the automated QIAamp Virus BioRobot MDx extraction platform (Qiagen, Hilden, Germany). V1–V4 *env* amplicons were generated by one step RT-PCR using primers Env-6210F (nt 6221–6245 on JR-CSF, Genbank accession number M38429) and Env-R3 (nt 7507–7527 on JR-CSF) and SuperScript™ III RT/Platinum® *Taq* High Fidelity (Invitrogen, Paisley, United Kingdom). Reactions were performed in sevenfold and amplicons were pooled and agarose gel purified (QiaQuick Gel Extraction kit, Qiagen). V1–V4 *env* amplicons were cloned in pCR4-TOPO vector and transformed into competent TOP10 *E. coli* cells according to the manufacturer's recommendations (TOPO TA Cloning® Kit, Invitrogen). Individual colonies were picked for further analysis using the QpExpression robot (Genetix Limited, Hampshire, United Kingdom). V1–V4-containing TOPO plasmids were amplified by rolling circle amplification using the illustra TempliPhi DNA amplification kit (GE Healthcare, Vilvoorde, Belgium) according to the manufacturer's instructions. These amplicons served as a template for sequencing. Sequencing reactions were prepared using the BigDye® Terminator Cycle Sequencing kit v. 3.1 (Applied Biosystems, Foster City, California, USA) with primers T3 (5′-AATTAACCCTCACTAAAGGG-3′) and T7 (5′-TAATACGACTCACTATAGGG-3′). Sequencing products were run on an ABI3730xl automated sequencer. Sequence editing and contig assembly were performed using SeqScape v2.5 (Applied Biosystems Incorporated, Foster City, CA).

Sequence analysis

Nucleotide sequences were assembled using the BioEdit package (Hall, 1999). Sequences were aligned using Clustal W with manual correction (Thompson et al., 1994). Nucleotide gaps were assigned after amino acid conversion to maintain translation integrity.

The best fitting nucleotide-substitution model was selected according to the Aikaike Information Criterium (AIC) using Modeltest v3.7 (Posada and Crandall, 1998). Phylogenetic trees were reconstructed in PAUP* v4.0b10 (Phylogenetic Analysis Using Parsimony (*and other methods), version 4, Sinauer Associates, Sunderland, Massachusetts, http://paup.csit.fsu.edu/), starting from a stepwise-addition tree with random addition order under a heuristic maximum likelihood search that implemented tree bisection and reconnection (TBR). Bootstrap analysis was performed using the above mentioned conditions on 100 replicates. The tree in Fig. 1 was constructed based on the 300-bp C2–V3 sequence using the sequence of patient P 9539, infected with a subtype A virus, as outgroup. The trees in Fig. 3 were constructed using the V1–V4 sequences (approximately 1014 bp) obtained after limiting dilution sequencing or clonal sequencing. The trees were rooted using the baseline consensus sequence as outgroup. Tree diagrams were plotted with Treeview (Page, 1996).

Based on the amino acid sequence of the V3 loop, the PSSM bioinformatics program (http://mullinslab.microbiol.washington.edu/computing/pssm/) predicted CCR5 use for all viral strains isolated at baseline (Jensen et al., 2003). For the samples collected at later time points minor variants with a predicted CXCR4 tropism were seen only in the last two samples collected from patient B.

Statistical analysis

The *t* test was used for the comparison of mean pairwise genetic distances.

The ANOVA test was used for the comparison of regression lines. For all tests *p*-values of <0.05 were considered statistically significant. Statistical analysis was performed with StatistiXL 1.8.

Nucleotide sequences

The sequences obtained during this study received Genbank accession numbers genbank:EU008093 to genbank:EU008328 and genbank:EU846607 to genbank:EU846981.

Acknowledgments

The authors would like to thank Maté Ongenaert and Annelies Haegeman from the Department of Molecular Biotechnology of the Faculty of Bioscience Engineering, UGent for help with the phylogenetic analysis and Veerle van Eygen and Jacqueline Reynaerts for technical assistance.

References

Ball, S.C., Abraha, A., Collins, K.R., Marozsan, A.J., Baird, H., Quinones-Mateu, M.E., Penn-Nicholson, A., Murray, M., Richard, N., Lobritz, M., Zimmerman, P.A., Kawamura, T., Blauvelt, A., Arts, E.J., 2003. Comparing the ex vivo fitness of CCR5-tropic human immunodeficiency virus type 1 isolates of subtypes B and C. J. Virol. 77 (2), 1021–1038.

Carmichael, A., Jin, X., Sissons, P., 1996. Analysis of the human env-specific cytotoxic T-lymphocyte (CTL) response in natural human immunodeficiency virus type 1 infection: low prevalence of broadly cross-reactive env-specific CTL. J. Virol. 70 (12), 8468–8476.

Deacon, N.J., Tsykin, A., Solomon, A., Smith, K., Ludfordmenting, M., Hooker, D.J., Mcphee, D.A., Greenway, A.L., Ellett, A., Chatfield, C., Lawson, V.A., Crowe, S., Maerz, A., Sonza, S., Learmont, J., Sullivan, J.S., Cunningham, A., Dwyer, D., Dowton, D., Mills, J., 1995. Genomic structure of an attenuated quasi-species of HIV-1 from a blood-transfusion donor and recipients. Science 270 (5238), 988–991.

Draenert, R., Allen, T.M., Liu, Y., Wrin, T., Chappey, C., Verrill, C.L., Sirera, G., Eldridge, R.L., Lahaie, M.P., Ruiz, L., Clotet, B., Petropoulos, C.J., Walker, B.D., Martinez-Picado, J., 2006. Constraints on HIV-1 evolution and immunodominance revealed in monozygotic adult twins infected with the same virus. J. Exp. Med. 203 (3), 529–539.

Essajee, S.M., Kim, M., Gonzalez, C., Rigaud, M., Kaul, A., Chandwani, S., Hoover, W., Lawrence, R., Spiegel, H., Pollack, H., Krasinski, K., Borkowsky, W., 1999. Immunologic and virologic responses to HAART in severely immunocompromised HIV-1-infected children. Aids 13 (18), 2523–2532.

Ganeshan, S., Dickover, R.E., Korber, B.T.M., Bryson, Y.J., Wolinsky, S.M., 1997. Human immunodeficiency virus type 1 genetic evolution in children with different rates of development of disease. J. Virol. 71 (1), 663–677.

Goulder, P.J.R., Watkins, D.I., 2004. HIV and SIVCTL escape: implications for vaccine design. Nat. Rev. Immunol. 4 (8), 630–640.

Gunthard, H.F., Frost, S.D.W., Leigh-Brown, A.J., Ignacio, C.C., Kee, K., Perelson, A.S., Spina, C.A., Havlir, D.V., Hezareh, M., Looney, D.J., Richman, D.D., Wong, J.K., 1999. Evolution of envelope sequences of human immunodeficiency virus type 1 in cellular reservoirs in the setting of potent antiviral therapy. J. Virol. 73 (11), 9404–9412.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 98/98/NT. Nucl. Acids Symp. Ser. 41, 95–98.

Husman, A.M.D., Koot, M., Cornelissen, M., Keet, I.P.M., Brouwer, M., Broersen, S.M., Bakker, M., Roos, M.T.L., Prins, M., deWolf, F., Coutinho, R.A., Miedema, F., Goudsmit, J., Schuitemaker, H., 1997. Association between CCR5 genotype and the clinical course of HIV-1 infection. Ann. Intern. Med. 127 (10), 882.

Hutto, C., Zhou, Y., He, J., Geffin, R., Hill, M., Scott, W., Wood, C., 1996. Longitudinal studies of viral sequence, viral phenotype, and immunologic parameters of human immunodeficiency virus type 1 infection in perinatally infected twins with discordant disease courses. J. Virol. 70 (6), 3589–3598.

Ioannidis, J.P.A., Rosenberg, P.S., Goedert, J.J., Ashton, L.J., Benfield, T.L., Buchbinder, S.P., Coutinho, R.A., Eugen-Olsen, J., Gallart, T., Katzenstein, T.L., Kostrikis, L.G., Kuipers, H., Louie, L.G., Mallal, S.A., Margolick, J.B., Martinez, O.P., Meyer, L., Michael, N.L., Operskalski, E., Pantaleo, G., Rizzardi, G.P., Schuitemaker, H., Sheppard, H.W., Stewart, G.J., Theodorou, I.D., Ullum, H., Vicenzi, E., Vlahov, D., Wilkinson, D., Workman, C., Zagury, J.F., O'Brien, T.R., Genetics, I.M.A.H.H., 2001. Effects of CCR5-Delta 32, CCR2-641, and SDF-1 3' A alleles on HIV-1 disease progression: an international meta-analysis of individual-patient data. Ann. Intern. Med. 135 (9), 782–795.

Jensen, M.A., Li, F.S., van't Wout, A.B., Nickle, D.C., Shriner, D., He, H.X., McLaughlin, S., Shankarappa, R., Margolick, J.B., Mullins, J.I., 2003. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. J. Virol. 77 (24), 13376–13388.

Kirchhoff, F., Greenough, T.C., Brettler, D.B., Sullivan, J.L., Desrosiers, R.C., 1995. Brief report — absence of intact Nef sequences in a long-term survivor with nonprogressive Hiv-1 infection. N. Engl. J. Med. 332 (4), 228–232.

Liu, S.L., Schacker, T., Musey, L., Shriner, D., McElrath, M.J., Corey, L., Mullins, J.I., 1997. Divergent patterns of progression to AIDS after infection from the same source: human immunodeficiency virus type 1 evolution and antiviral responses. J. Virol. 71 (6), 4284–4295.

Lukashov, V.V., Kuiken, C.L., Goudsmit, J., 1995. Intrahost human-immunodeficiency-virus Type-1 evolution is related to length of the immunocompetent period. J. Virol. 69 (11), 6911–6916.

Mani, I., Gilbert, P., Sankale, J.L., Eisen, G., Mboup, S., Kanki, P.J., 2002. Intrapatient diversity and its correlation with viral setpoint in human immunodeficiency virus type 1 CRF02_A/G-IbNG infection. J. Virol. 76 (21), 10745–10755.

Markham, R.B., Wang, W.C., Weisstein, A.E., Wang, Z., Munoz, A., Templeton, A., Margolick, J., Vlahov, D., Quinn, T., Farzadegan, H., Yu, X.F., 1998. Patterns of HIV-1 evolution in individuals with differing rates of CD4 T cell decline. Proc. Natl. Acad. Sci. U. S. A. 95 (21), 12568–12573.

Marozsan, A.J., Torre, V.S., Johnson, M., Ball, S.C., Cross, J.V., Templeton, D.J., Quinones-Mateu, M.E., Offord, R.E., Arts, E.J., 2001. Mechanisms involved in stimulation of human immunodeficiency virus type 1 replication by aminooxypentane RANTES. J. Virol. 75 (18), 8624–8638.

Mikhail, M., Wang, B., Lemey, P., Beckholdt, B., Vandamme, A.M., Gill, M.J., Saksena, N.K., 2005. Full-length HIV type 1 genome analysis showing evidence for HIV type 1 transmission from a nonprogressor to two recipients who progressed to AIDS. AIDS Res. Hum. Retrovir. 21 (6), 575–579.

Page, R.D.M., 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12 (4), 357–358.

Parren, P.W.H.I., Moore, J.P., Burton, D.R., Sattentau, Q.J., 1999. The neutralizing antibody response to HIV-1: viral evasion and escape from humoral immunity. Aids 13, S137–S162.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14 (9), 817–818.

Rangel, H.R., Weber, J., Chakraborty, B., Gutierrez, A., Marotta, M.L., Mirza, M., Kiser, P., Martinez, M.A., Este, J.A., Quinones-Mateu, M.E., 2003. Role of the human immunodeficiency virus type 1 envelope gene in viral fitness. J. Virol. 77 (16), 9069–9073.

Rosenberg, E.S., Altfeld, M., Poon, S.H., Phillips, M.N., Wilkes, B.M., Eldridge, R.L., Robbins, G.K., D'Aquila, R.T., Goulder, P.J.R., Walker, B.D., 2000. Immune control of HIV-1 after early treatment of acute infection. Nature 407 (6803), 523–526.

Shankarappa, R., Margolick, J.B., Gange, S.J., Rodrigo, A.G., Upchurch, D., Farzadegan, H., Gupta, P., Rinaldo, C.R., Learn, G.H., He, X., Huang, X.L., Mullins, J.I., 1999. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. J. Virol. 73 (12), 10489–10502.

Strunnikova, N., Ray, S.C., Livingston, R.A., Rubalcaba, E., Viscidi, R.P., 1995. Convergent evolution within the V3 loop domain of human-immunodeficiency-virus type-1 in association with disease progression. J. Virol. 69 (12), 7548–7558.

Setthent, R., Foongladda, S., Chearskul, S., Wanprapa, N., Likanonskul, S., Kositanont, U., Riengrojpitak, S., Sahaphong, S., Wasi, C., 1998. V3 sequence diversity of HIV-1 subtype E in infected mothers and their infants. J. Acquir. Immune Defic. Syndr. 18 (4), 323–331.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. Improved sensitivity of profile searches through the use of sequence weights and gap excision. Comput. Appl. Biosci. 10 (1), 19–29.

Troyer, R.M., Collins, K.R., Abraha, A., Fraundorf, E., Moore, D.M., Krizan, R.W., Toossi, Z., Colebunders, R.L., Jensen, M.A., Mullins, J.I., Vanham, G., Arts, E.J., 2005. Changes in human immunodeficiency virus type 1 fitness and genetic diversity during disease progression. J. Virol. 79 (14), 9006–9018.

Wolinsky, S.M., Korber, B.T.M., Neumann, A.U., Daniels, M., Kunstman, K.J., Whetsell, A.J., Furtado, M.R., Cao, Y.Z., Ho, D.D., Safrit, J.T., Koup, R.A., 1996. Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. Science 272 (5261), 537–542.

Zhu, T.F., Mo, H.M., Wang, N., Nam, D.S., Cao, Y.Z., Koup, R.A., Ho, D.D., 1993. Genotypic and phenotypic characterization of HIV-1 in patients with primary infection. Science 261 (5125), 1179–1181.